

Procesamiento de Información Temporal

Tema 1: Modelos Clásicos de Series Temporales Parte 1: Introducción y Modelo AR

Alicia Lozano Díez

alicia.lozano@uam.es

Audias – Audio, Data Intelligence and Speech

Universidad Autónoma de Madrid

<http://audias.ii.uam.es>

Materiales basados en los de Daniel Ramos Castro

1

Contenidos: Tema 1 Parte 1

- Introducción a las Series Temporales
- Modelos de Filtrado Lineal
- Ejemplo práctico: Modelo AR para codificación de voz
(Coeficientes de Predicción Lineal, LPC)

Introducción a las Series Temporales

3

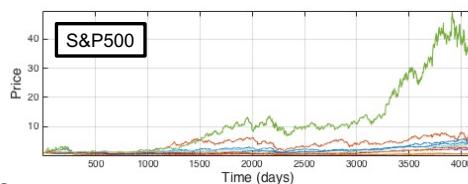
Introducción a las Series Temporales

- Serie Temporal (también conocida como señal en tiempo discreto)
 - Secuencia de observaciones (datos) tomadas periódicamente en el tiempo
- Aparecen en miles de situaciones
 - Economía, finanzas, ingeniería, ciencias naturales, ciencias sociales...
- Áreas de aplicación
 - Predicción de valores futuros
 - Análisis para conocer el comportamiento
 - Caracterización de un sistema (función de transferencia)
 - Extracción de información
 - ...

4

Series Temporales: Ejemplos

- Valores diarios de activos financieros



- Áreas de interés

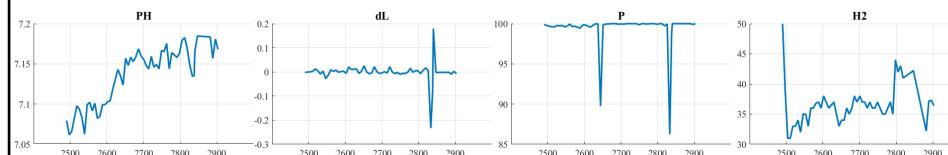
- Predecir valores futuros para inversión
 - Analizar comportamientos y tendencias
 - Generación de series sintéticas
 - Realistas:



Javier Franco-Pedroso, Joaquín González-Rodríguez et al. "Generating virtual scenarios of multivariate financial data for quantitative trading applications." The Journal of Financial Data Science, 2019.

Series Temporales: Ejemplos

- Mediciones de sensores e instrumentos industriales



Daniel Ramos, Pablo Ramírez et al. "Dynamic Bayesian Networks for Temporal Prediction of Chemical radioisotope Levels in Nuclear Power Plant Reactors". Chemometrics and Intelligent Laboratory Systems, 2021.

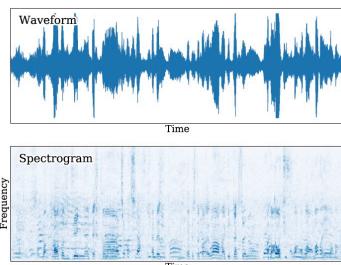
- Áreas de interés

- Predecir anomalías
 - Mantenimiento preventivo
 - Predecir valores de variables deseadas
 - Rendimiento de una central eléctrica, por ejemplo
 - Predecir efectos de una política de operación de la planta

Series Temporales: Ejemplos

■ Señales de audio y voz

Diego de Benito, Alicia Lozano-Díez et al. "Exploring convolutional, recurrent, and hybrid deep neural networks for speech and music detection in a large audio dataset". EURASIP Journal on Audio, Speech and Music Processing, 2019.



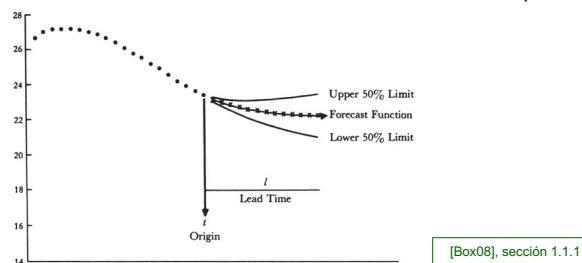
■ Áreas de interés

- Detección de eventos (timbre, televisión, música, perros...)
- Reconocimiento de voz (mensaje hablado) y palabras clave
- Reconocimiento de idioma
- Diarización y reconocimiento de locutores

Series Temporales: Aplicaciones

■ Predicción

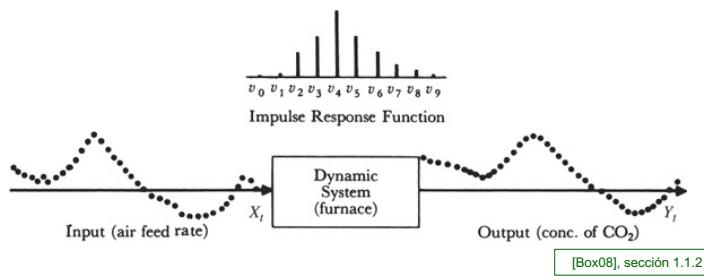
- Análisis de una serie temporal en el pasado
- Para predecir el futuro
- En un tiempo de interés (lead time)
- Con unos intervalos de credibilidad (probabilidad del X%)
- Minimizando los errores con respecto a la serie “real” (de entrenamiento)
- Evaluando el error con valores futuros de una serie “real” (de test)



[Box08], sección 1.1.1

Series Temporales: Aplicaciones

- Caracterización de un sistema
 - Entradas vs. salidas: sistema dinámico
 - Si el sistema es lineal e invariante en el tiempo: filtro
 - Se puede estimar la función de transferencia o la respuesta al impulso/en frecuencia de dicho filtro
 - O se puede establecer un *modelo* del filtro que nos permita hacer predicción o caracterizar el sistema



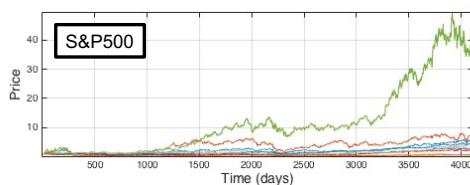
Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales

9



Series Temporales: Aplicaciones

- Series multivariadas
 - Idea: varias series ocurriendo al mismo tiempo



- Unos valores de la serie *influyen* (es decir, *dan información*) sobre otros
 - En diferentes instantes temporales
 - En el mismo instante temporal
- No caracterizar la relación entre esos valores implica...
 - Perder esa valiosa información
 - Para predicción, para análisis, etc.

Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales

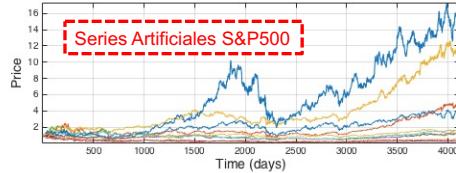
10



10

Tipos de Modelos de Series Temporales

- Deterministas
 - Normalmente basados en sistemas dinámicos
 - Explotan la física del mecanismo de generación de la serie
 - Ecuación diferencial o ecuación en diferencias
- Estocásticos
 - Series son realizaciones de un proceso estocástico
 - Modelado de la serie
 - Intrínsecamente busca modelar dicho proceso
 - Clave las suposiciones sobre estacionariedad, estacionalidad, eventos no relacionados, etc.



Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales

11



11

Modelos Estocásticos

- Modelos estacionarios
 - Suponen algún tipo de estacionariedad
 - Típicamente estacionariedad débil (o en sentido amplio)
 - Más homocedasticidad
 - Modelos AR, MA, ARMA
- Modelos no estacionarios
 - Suponen que la media del proceso no es constante con el tiempo
 - Modelos ARIMA
 - Suponen que existe estacionalidad (perturbación pseudo-periódica)
 - Modelos SARIMA
- Influencia de variables exógenas
 - Series temporales independientes de los valores de interés
 - Que son parte de la serie bajo análisis
 - SARIMAX

Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales

12



12

Notación

- [Box08] utiliza operadores
 - Sea z_t el valor de la serie en el instante t :
 - Operador “retardo” (backward shift)
 - $z_{t-m} = B^m z_t$
 - Operador “adelanto” (forward shift)
 - $z_{t+m} = F^m z_t = B^{-m} z_t$
 - Operador diferencia o gradiente discreto
 - $\nabla z_t = z_t - z_{t-1} = (1 - B)z_t$

Notación

- Los operadores retardo y adelanto de [Box08] se relacionan con la transformada Z
 - Poderosa herramienta de análisis (generaliza la T. Fourier en t. discreto)
 - Permite ver ciertos modelos como un “filtro” (sistema lineal invariante en el tiempo)
 - En este caso, filtro basado en ecuaciones en diferencias
- De este modo:
 - Sea z_t el valor de la serie en el instante t :
 - Operador “retardo” (backward shift) es, en transformada Z:
 - $z_{t-m} = B^m z_t = Z^{-m} z_t$
 - Operador “adelanto” (forward shift)
 - $z_{t+m} = F^m z_t = B^{-m} z_t = Z^m z_t$
 - Operador diferencia o gradiente discreto
 - $\nabla z_t = z_t - z_{t-1} = (1 - Z^{-1})z_t$

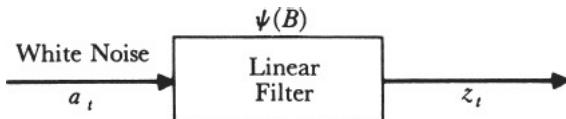
Notación

- ¡Ojo! Cuidado con las notaciones...
 - Notación [Box08]:
 - Llama z_t a los valores de la serie en el instante de tiempo (discreto) t
 - Utiliza operadores B y F
 - Para transformada Z, la notación de [Oppenheim10]:
 - Llama $x[n]$ a los valores de la serie (señal x) en el instante de tiempo discreto n
 - Utiliza operadores Z^{-1} y Z
 - Dependiendo del texto al que acudamos, usaremos una u otra notación

Modelos Basados en Filtros Lineales

Modelo de Filtro Lineal

- Una serie temporal puede ser generada por
 - Una secuencia típicamente de “ruido blanco” $a_t \sim \mathcal{N}(0, \sigma_a^2)$
 - Cada coeficiente (“shock”) está incorrelado con resto (ruido blanco)
 - Serie a_t no observable, serie z_t observada
 - Un filtro lineal que transforma esa secuencia
- Muchas series temporales se pueden analizar bien con este modelo



$$z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots$$

[Box08]

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$$

Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales

17



Modelo de Filtro Lineal

- Si el filtro lineal es finito (FIR), o infinito y absolutamente sumable (IIR)...

$$\sum_{j=0}^{\infty} |\psi_j| < \infty$$

- Filtro estable
 - Serie z_t estacionaria
- Y viceversa
 - Filtros no estables dan lugar a series no estacionarias
 - Por ejemplo, series que “crecen” hasta el infinito
 - La media del proceso subyacente aumenta con el tiempo: no estacionaria

Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales

18



Modelo Autorregresivo (AR)

- El filtro lineal es sobre valores pasados de la señal
 - “Centrados” (se resta μ) $\tilde{z}_t = z_t - \mu$
 - Con un ruido aleatorio independiente o “shock” (a_t)
- Modelo autorregresivo de orden p

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \cdots + \phi_p \tilde{z}_{t-p} + a_t$$

- Regresión lineal con los propios valores pasados: “Autorregresión”

- Operador autorregresivo:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad \phi(B)\tilde{z}_t = a_t$$

- Equivalente a un filtro en transformada Z

Modelo Autorregresivo (AR)

- Problema:
 - Encontrar los valores de los coeficientes del operador AR (filtro en transformada Z)
 - Además del offset μ y de la varianza del ruido $a_t \sim \mathcal{N}(0, \sigma_a^2)$
 - Problema de encontrar $p+2$ parámetros
- Estabilidad del filtro y estacionariedad de la serie modelada
 - Un proceso AR es un caso particular de un filtro lineal

$$\phi(B)\tilde{z}_t = a_t \quad \tilde{z}_t = \phi^{-1}(B)a_t = \psi(B)a_t$$

- Estable si y solo si $\sum_{j=0}^{\infty} |\psi_j| < \infty$
- Equivalente a decir que el filtro definido por la transformada Z es estable

Ejemplo de Modelo AR Análisis de Predicción Lineal Para Codificación de Voz

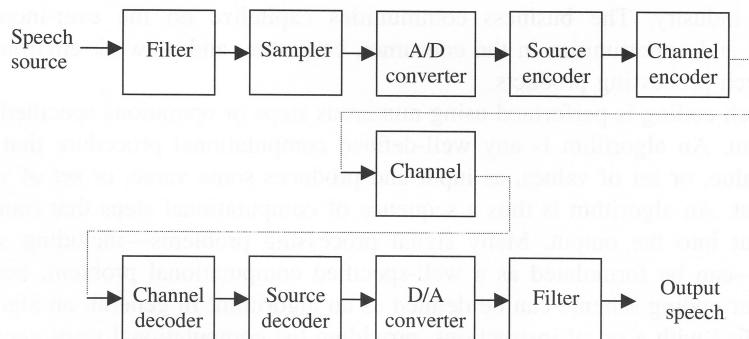
21

Modelo AR: Predicción Lineal (LPC)

- Aplicación de un modelo AR
 - Para codificación de audio y voz sin pérdidas
- Problema:
 - Tenemos una señal de audio o voz
 - Queremos reducir significativamente su tamaño
 - MP3: codificación con pérdidas (audio original diferente a audio comprimido)
 - FLAC: codificador sin pérdidas (audio original prácticamente igual al comprimido)
- FLAC basado en un modelo AR
 - Coeficientes de Predicción Lineal (LPC)
 - En realidad es mucho más complejo, pero se basa en LPC
- La hipótesis es lo que vale para la señal de voz, también valdrá para la señal de audio
 - No del todo cierta, pero funciona moderadamente bien

22

Diagrama de bloques de un codificador genérico de audio o voz



Reproducción del libro: Speech Coding Algorithms, Wai C. Chu, John Wiley & Sons

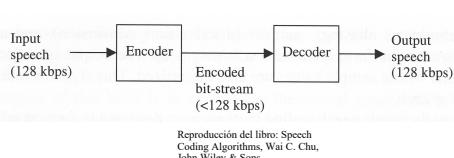
La señal digital de audio: La señal de CD

- Un ejemplo prototípico de señal digital de audio es la señal de CD
- Se caracteriza por:
 - Frecuencia de muestreo: 44.1 kHz
 - Bits/muestra: 16 bits
 - Canales: 2 (estéreo)
 - En algunos casos (sonido 3D) emplearemos más canales
- Con esto la tasa binaria de fuente que se obtiene es de:
 - $R = 44100 \times 16 \times 2 = 1411200 \text{ bits/segundo} = 1.4112 \text{ Mb/segundo}$.
 - 3 minutos de audio: $(180s \times 1,4112 \text{ Mb/s})/(8\text{bit}/\text{byte})=30,28 \text{ MB (!!)}$
- El objetivo de la codificación de audio es reducir drásticamente esta tasa binaria sin afectar demasiado a la calidad del audio.
 - **En este ejemplo, se hace sin afectar en absoluto la calidad del audio: sin perdidas**

La señal digital de voz

- La señal digital de voz típicamente se caracteriza por:
 - Frecuencia de muestreo: 8 kHz.
 - Bits/muestra: 16 bits
 - Canales: 1 (mono)
- Con esto la tasa binaria de fuente que se obtiene es de:
 - $I = 8000 \times 16 \times 1 = 128000$ bits/segundo = 128 Kb/segundo
 - Canal de transmisión GSM FR (full rate): 13 Kb/segundo
- El objetivo de la codificación de voz es reducir drásticamente esta tasa binaria sin afectar demasiado a la calidad del audio
 - **En este ejemplo lo haremos sin afectar en absoluto la calidad del audio: sin perdidas**

Diagrama de bloques simplificado de un codificador de voz / audio

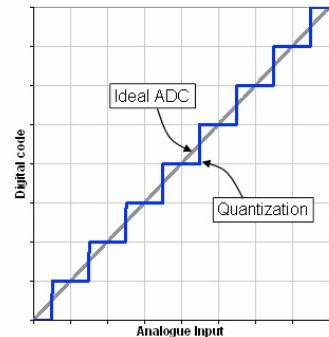
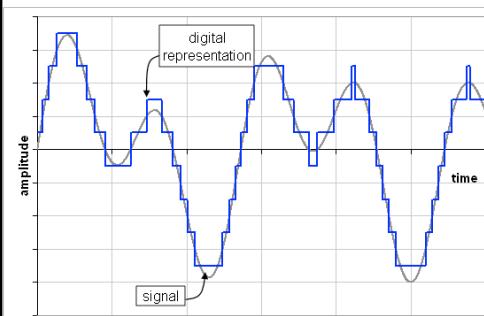


Reproducción del libro: Speech Coding Algorithms, Wai C. Chu, John Wiley & Sons

- Nos centraremos en la pareja codificador de fuente / decodificador de fuente, que:
 - Transforman la señal digital de audio / voz en una señal con menor tasa binaria
 - Transforman de nuevo esta señal en una señal digital de audio / voz similar a la original
- De esta forma, emplearemos el esquema simplificado de la figura

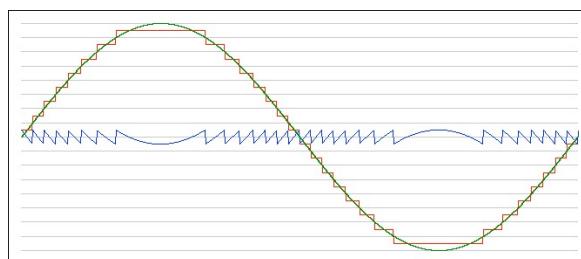
Fundamentos: Cuantificación

- Cuantificar una serie temporal de valores **continuos** es asignarle niveles discretos de amplitud



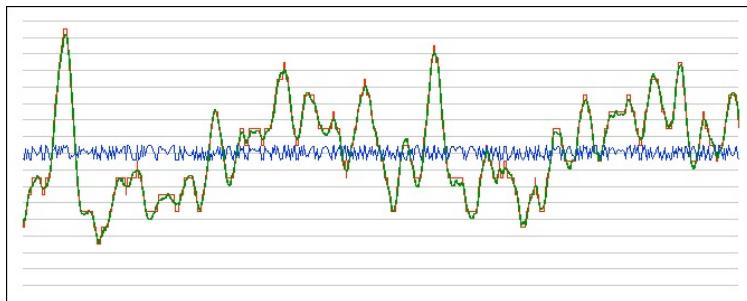
Fundamentos: Ruido de Cuantificación

- Cuando se cuantifica una señal continua, siempre se introduce error
 - “Error de cuantificación”, señal cuantificada menos la señal original
- La señal original (verde) más el error de cuantificación (azul) es igual a la señal cuantificada (roja)
 - Se puede tomar como un error aditivo \Rightarrow ruido aditivo
 - La señal cuantificada es “degradada” por el ruido aditivo)
 - También llamado “ruido de cuantificación”



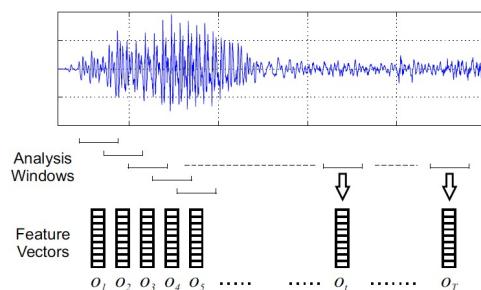
Fundamentos: Ruido de Cuantificación

- Ruido de cuantificación en señal de audio real

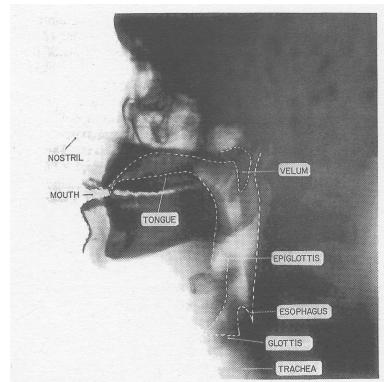


Fundamentos: Análisis a Corto Plazo

- Análisis de voz/audio a corto plazo: primero se segmenta la voz en tramas cortas
 - Típicamente entre 20 y 100 ms.
 - Típicamente solapadas entre ellas un porcentaje, por ejemplo el 50%
 - En esas tramas cortas hay dos tipos de sonidos
 - Sonidos "sordos": ruido estacionario
 - Sonidos "sonoros": pseudo-periódicos
- De cada trama a corto plazo, se halla una transformada
 - Típicamente Dicrete (Fast) Fourier Transform (DFT/FFT)



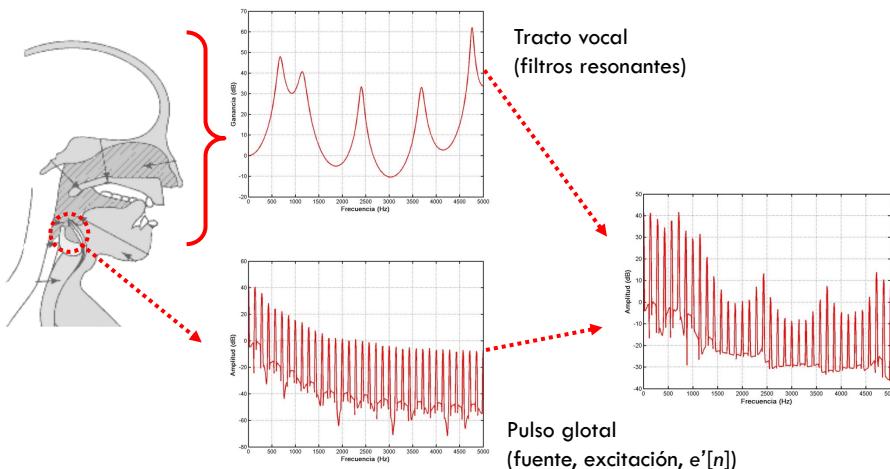
El mecanismo de producción de voz



Reproducción del libro
"Digital Processing of Speech Signals",
L. R. Rabiner, R. W. Schafer,
Prentice-Hall.

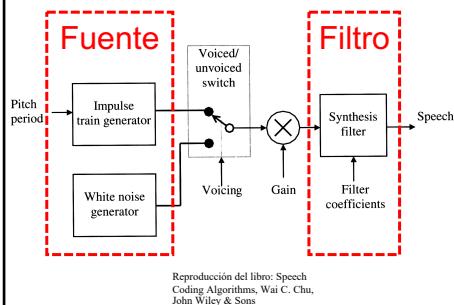
- La figura representa una radiografía que muestra los órganos más importantes del aparato fonador humano:
 - El *Tracto Vocal* (línea punteada): comienza en las cuerdas vocales (glotis) y termina en los labios
 - El tracto vocal consiste en:
 - Faringe (conexión desde la tráquea a la boca).
 - Cavidad oral
 - El tracto vocal mide:
 - Entre 14 y 17 cm de longitud en media
 - De 0 (cierre completo) a 20cm² de sección transversal
 - El *Tracto Nasal* (no se aprecia demasiado en la figura): comienza en el velo y termina en la nariz
 - El velo puede bajar conectando acústicamente el tracto nasal con el tracto vocal para producir sonidos nasales

Fundamentos: Modelo Simplificado del Tracto Vocal: Fuente y filtro



Codificación de voz por predicción lineal (Linear Prediction Coding, LPC)

- Modelo simplificado de producción de voz (fuente y filtro)
- ¿Cómo usar este modelo para codificación de voz?
 - Asumiendo que podemos separar la señal de voz en segmentos (*frames*, tramas, "ventanas") suficientemente cortos como para que los parámetros del modelo no varíen
 - Estimando los parámetros del modelo para cada uno de esos *frames*, de derecha a izquierda según la figura, porque partimos de la señal de voz y vamos analizándola en esa dirección:
 - Coeficientes del filtro de síntesis
 - Ganancia: dependiendo de la energía de la excitación
 - Sonoridad: sordo o sonoro
 - Período fundamental (*pitch period*): período de la señal de excitación en sonidos sonoros



Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales

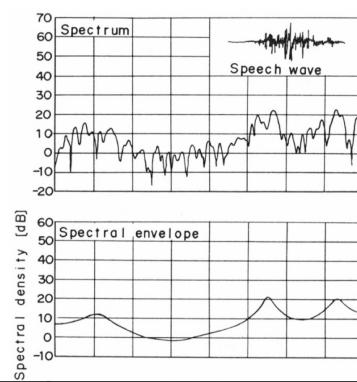
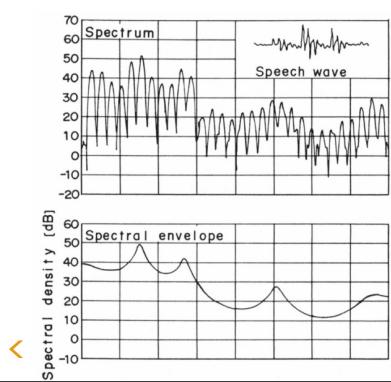


33

33

El tracto vocal como filtro: formantes

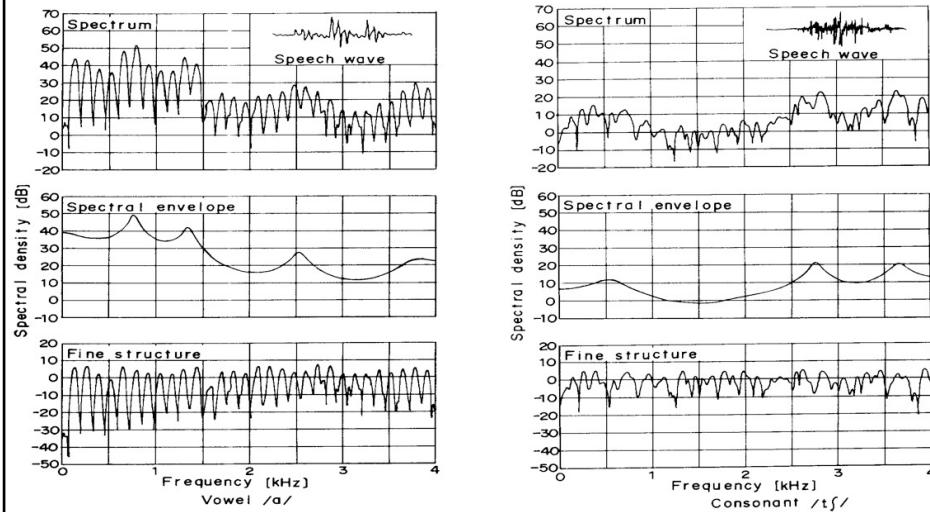
- Las cavidades oral, faríngea y nasal del tracto vocal forman un sistema de resonadores
- Cada resonancia del tracto vocal, llamada formante, está caracterizada por:
 - Frecuencia formántica (en Hz) y amplitud relativa (en dB)
 - Las frecuencias formantes reflejan el objetivo articulatorio del hablante
 - Ancho de banda (en Hz)
 - Reflejan pérdidas acústicas



34

17

Descomposición espectral: excitación (sordo/sonoro) + tracto vocal



35

Cálculo de los coeficientes del filtro de síntesis: Predicción lineal (Modelo AR)

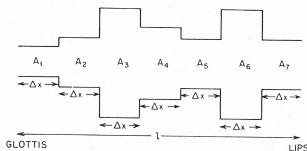
- La predicción lineal forma parte integral de casi todos los codificadores de voz modernos.
- La idea fundamental consiste en predecir una muestra de voz como una combinación lineal de muestras anteriores
 - Si denotamos la señal discreta de voz como $x[n]$, la predicción de voz trata de encontrar los coeficientes a_k que permiten expresar aproximadamente una muestra en función de P muestras anteriores (donde P es el orden de la predicción lineal)

$$x[n] \approx \sum_{k=1}^P a_k x[n-k] \Rightarrow x[n] = \sum_{k=1}^P a_k x[n-k] + e[n] = \sum_{k=1}^P a_k x[n-k] + G e'[n]$$
 - Obviamente, la predicción con un P pequeño no será perfecta, y cometeremos un error de predicción, $e[n]$, relacionado con la excitación teórica $e[n]$ por una ganancia G
 - Para cada frame de voz los pesos empleados para la combinación lineal se eligen de forma que se minimice el error cuadrático medio (energía media) del error de predicción $e[n]$
 - Los pesos resultantes a_k se denominan *coeficientes de predicción lineal*

36

Relación entre la predicción lineal y el modelo de producción de voz (1)

- ¿Cuál es la justificación de la predicción lineal? ¿Por qué funciona tan bien para codificación de voz?
- La justificación para la predicción lineal se basa en el modelo digital del tracto vocal, $V(z)$, basado en una concatenación de secciones de tubos acústicos.



Reproducción del libro
"Digital Processing of Speech Signals",
L. R. Rabiner, R. W. Schafer,
Prentice-Hall.

- Analizando este modelo, llegábamos a la siguiente función de transferencia:

$$V(z) = \frac{X(z)}{E(z)} = \frac{G}{1 - \sum_{k=1}^N a_k z^{-k}} = GH(z) \quad H(z) = \frac{X(z)}{E(z)} = \frac{1}{1 - \sum_{k=1}^N a_k z^{-k}}$$

- Donde N era el número de secciones de tubos acústicos empleados en el modelo

Relación entre la predicción lineal y el modelo de producción de voz (2)

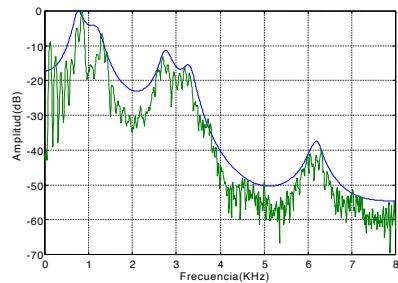
- Esta función de transferencia define la siguiente relación entre la entrada $e[n]$ y la salida $x[n]$ de la señal acústica en el tracto vocal:

$$V(z) = \frac{E'(z)}{X(z)} = \frac{1 - \sum_{k=1}^N a_k z^{-k}}{G} \Rightarrow x[n] = \sum_{k=1}^N a_k x[n-k] + Ge'[n]$$

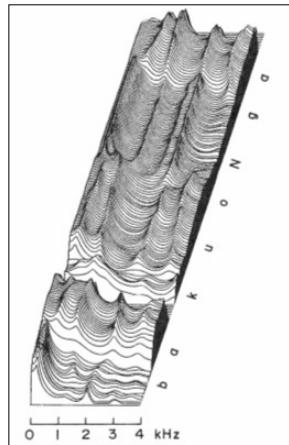
- Como vemos, la función de transferencia del modelo digital del tracto vocal, $V(z)$, que obtuvimos al analizar el mecanismo de producción de voz nos está diciendo que cada muestra de salida de la señal de voz depende linealmente de las N muestras anteriores (siendo N el número de secciones consideradas en el modelo del tracto vocal), así como de la excitación multiplicada por una ganancia.
 - Los sistemas en los que la salida depende sólo de salidas en instantes anteriores y de la entrada actual se denominan sistemas autorregresivos (AR).
 - También se denominan autorregresivas (AR) las salidas de dichos sistemas.
- La coincidencia con la fórmula de predicción lineal es evidente haciendo que:
 - El orden de predicción lineal, P , coincide con el número de secciones transversales empleadas en el modelo de producción de voz, N
 - El error de predicción $e[n]$, sea igual a la ganancia G por la excitación del tracto vocal, $e'[n]$
 - $x[n]$ sea la señal acústica a la salida del tracto vocal

Filtro de predicción lineal

- Espectro de la señal de voz (verde).
- Filtro de predicción lineal obtenido a partir de la señal de voz (azul).



$$V(z) = \frac{X(z)}{E'(z)} = \frac{G}{1 - \sum_{k=1}^N a_k z^{-k}} = GH(z)$$



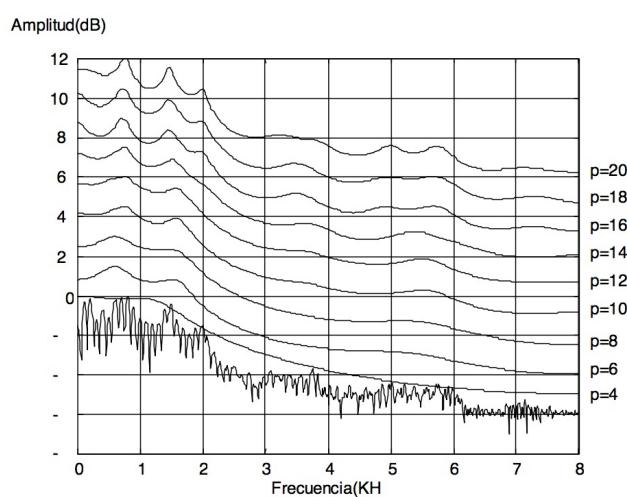
Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales



39

39

Ajuste del orden de predicción p



Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 1: Modelos Clásicos de Análisis de Series Temporales



40

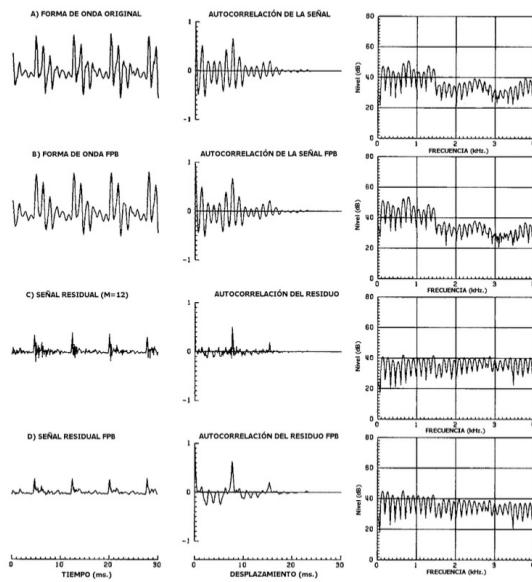
40

20

Señal residual del análisis LPC

- El análisis LPC permite la extracción de la señal residual de una trama de voz
- Dicho residuo contiene la estructura fina espectral "blanqueada" (sin el efecto del tracto vocal)
- En codificación de voz, se usa para excitar el filtro vocal definido por los p coeficientes LPC
- También se usa para estimación de pitch, calculando la autocorrelación de la señal residual

 Procesamiento de Información Temporal. MÁSTER EN CIENCIA DE DATOS. Tema 1: Modelos Clásicos de Análisis de Series Temporales

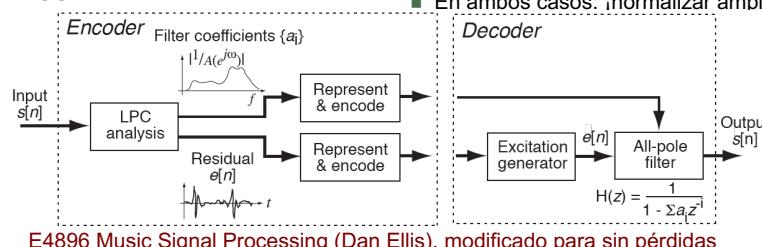


41

Estrategia de Compresión Sin Pérdidas

- Codificador:
 - Análisis a corto plazo y enventanado
 - Se obtiene $s[n]$, trama enventanada
 - Análisis LPC. Se obtienen:
 - Coeficientes del filtro LPC, a_k
 - Señal error $e[n]$
 - Almacenar en el fichero, para cada $s[n]...$
 - Coeficientes del filtro LPC
 - Versión comprimida sin pérdidas de $e[n]$

- Decodificador:
 - Para cada ventana
 - Descomprimir $e[n]$ ("excitation generator")
 - Filtro de síntesis LPC a $e[n]$ para obtener $s[n]$
 - Suma de todas las tramas de audio
 - Si están solapadas: Overlap-add
 - Las ventanas solapadas sumadas deberían ser constantes...
 - Si no están solapadas: dividir por la ventana
 - En ambos casos: ¡normalizar amplitud!



Procesamiento de Información Temporal. MÁSTER EN CIENCIA DE DATOS
Tema 1: Modelos Clásicos de Análisis de Series Temporales





42

Referencias Bibliográficas

- [Box08] George E. P. Box, Gwilym M Jenkins and Gregory C Reinsel. "Time series analysis : forecasting and control, 4th ed.". John Wiley and Sons, 2008.
- [Oppenheim10] Alan V. Oppenheim, Ronald W. Schaffer, "Tratamiento de Señales en Tiempo Discreto, 3^a ed." Pearson Education, 2011.
- [Benesti08] Benesti et al. "Linear Prediction". In Handbook of Speech Processing, Chapter 7, Springer, 2008.